Peer Reviewed Article   Vol.1 No.(2/3) September 1999

Strategies for the hierarchical selection of information in a volatile information-rich environment

Erna Fourie and Tom Bisschoff
Rand Afrikaans University
hamf@edcur.rau.ac.zaContents

Contents

Introduction

Our proposal submitted to the conference committee started off with the following statement:

'Learning is, amongst other things, based on the availability and accessibility of information. Information is available; it is communicated and used by people around the world and in every sphere. Reference is made to the information highway, the Internet and computer mediated communication as terminology pertaining to the storage and availability of information.'

This paper is not the result of research by the presenters, but rather, reflects experiences and solutions to problems encountered in accessing the new and volatile world of information. This is not a 'model' but rather 'strategies'. The work is still in progress (or in the language of the Web – under construction) and will be carried on by doing interviews with a number of frequent users (frequent being people using the Internet for specific information searches on a daily basis) (compare Newby, Stepich, Lehman & Russell, 1996:134).

.   Top

The world of knowledge and information

The world of knowledge has been and is still exploding and ways need to be developed to structure this world according to the needs of the user of the available information. The needs of the user are expressed in the inquiry carried out.

Popper (Burke, 1983:105, 106; Keeves, 1996:1; O'Hear 1980:171; Popper, 1992:7, 9, 20,)

has distinguished between three different worlds involved in human inquiry. The entities of the real world, which include physical objects, as well as the various structures, created by human society and which include business and training, form world 1. World 2 is the world of subjective experiences, a world of individual mental states, that comprises the states of conscious thought and psychological dispositions as well as the unconscious states of mind of individuals. Wisdom is an entity in world 2. There is, however, in addition world 3 that has been created as a new objective world that is the product of human minds. It should be noted that world 3 not only contains the corporate body of pre-positional knowledge concerned with causal explanation, but also the works of art, music and literary writings that are all part of the world of shared knowledge. The important point is that world 3 objects have acquired a reality of their own. Popper emphasises that these three worlds, which are interconnected and act upon each other in some way, also partially overlap each other. The word world is obviously not being used here to mean the universe or cosmos but rather parts of it (Popper, 1992:8).

A forth world has, however, developed over the past two decades. This world 4 can also be described as the world of information.

World 4: the world of information

The world of information emanates from the first three worlds and is an ever changing and therefore a volatile world. This change is not measured in months but in seconds.

In the pre-Internet era changes that took place had to be disseminated according to traditional protocol, i.e. paper based. With the advent of the Internet and the related 'information highway' any changes in existing information can be integrated into the relevant field momentarily and sent into cyberspace.

Cyberspace can be intimidating to the novice venturing into it and without a strategy to access the information; it has no meaning for a user in need of information. This assumption is the one focus of this paper.

A self-generated curriculum in the business world

In the previous paragraph, the one focus of this paper was identified as a strategy to access information. In this section, attention is given to the reason why information (knowledge of tomorrow today) must be accessed by the 'learner' in the business world. 'Scholars have yet to examine fully the growing use of technology to let students take control of their own learning' (Kent & McNergney, 1998:5).

In education and training it is common practice to refer to a course, related outcomes and the curriculum. When these perspectives are outlined the relevant content can be selected. The first logical step that follows is the presentation of content by means of, for example, face-to-face teaching, correspondence or computer mediated delivery. Next are the mastery of the content and some form of evaluation or assessment (compare Oliva, 1992:482). This scenario has not changed much over the decades but the volatile and dynamic world of business and technology has increased the rate at which these steps follow each other exponentially.

To teach today the knowledge of tomorrow, the individual in the business environment

becomes a learner who must be guided to construct his or her own 'course' and related objectives or outcomes, curriculum and guidelines as to how to select content, rather than to outline the content from the 'teacher's' perspective. Electronic communication has made every learning situation an information-rich learning environment (compare Simonson & Thompson, 1997:226).

The central question in this paper could now be stated as follows: By what strategies could the individual learner select relevant information from an electronic source, such as the Internet in a personalised 'classroom', with personalised objectives or outcomes and a personalised curriculum (compare Pratt, 1994:65).

Information technology and education

1. **Access of information from world 4**

The massive volumes of information that exists in world 4 (remember it is not reality itself – world 1) and also not world 3 which is a structured body of knowledge. As soon as world 4 becomes part of world 3 it is not information any longer but is almost 'old news'.

To understand the process of access one has to start with world 2 (the learner's world). He or she experiences an uncomfortable situation because he/she is confronted with a reality (world 1) that he/she cannot handle. Or cannot handle as well as he/she would like to handle the situation.

A practical example: A Growth fund manager realises (I do not know!) that the millions being invested for shareholders are not yielding the return on investment. He/she is making the wrong decisions. He/she realises that the information available to him/her is not the information he/she needs (I need to know) to make the correct decisions. Now he/she needs to ask the question: What do I need to know?

2. **Designing the curriculum**

Mention has been made of the self-generated curriculum. Albeit a self-generated curriculum, the design principles are the same as that for the design of a curriculum for a formal course. The principles are, according to most authors: needs analysis; outcomes (objectives); content; learning experiences; and learning opportunities and evaluation (compare Stenhouse, 1975:53).

The question that must be attended to remains: 'What do I need to know?'

The learner reflects on his/her experiences (world 2), looks at the real world (world 1) and its impact on the situation. He/she reflects on the existing (clearly insufficient) body of knowledge (world 3) and establishes what he/she needs to know to solve the problem. What are the outcomes the learner wants to achieve?

With this reflective knowledge he/she can start utilising the world of information (World 4) from which the content must now be selected. Learning opportunities and experiences must now be found and assessment of whether the outcomes have been reached must be carried out.

In this paper we would like to give credit to the role of libraries, formal degrees, short courses, colleagues and mentors. However, bearing the perishability of knowledge in mind, a

source of information that draws attention is the Internet.

The Internet here not only represents a database but also becomes an electronic classroom. Somekh and Davis (1997:178) express it as follows: '…electronic communications can open up the classroom, or a whole school, to the world. It can be a means to enhance and extend curriculum and professional development on-site and over time. However, its use must be accompanied by increased responsibility of the learner and a consequent restructuring of the teaching (learning) process …'.

Hierarchical selection of information – strategies

1. **The Internet**

This massive world wide connection of many thousands of computer networks has developed over the past number of years and, although there is no one person who can claim ownership of the internet, there is a dynamic and incessant flow of information which is placed on the internet minute by minute.

It is, however, very time consuming and could be very unsuccessful if unstructured or unplanned searches are conducted. To deal with all these negative experiences, four strategies will now be suggested and illustrated.

2. **Strategies**

A strategy could be defined as the art of war or the management of an army. The Internet could therefore be attacked in the following ways:

- Problem-solving approach
- Chaos theory approac
- Pattern-recognition approach
- Search engine protocol approach

2.1 **Problem-solving approach**

'Problem solving activities are part of everyday life. Many of us, however, do not approach this process systematically, and as a result, solutions may often be hit or miss' (Forcier, 1996:110). It is not really possible to identify a 'best' approach. In this connection work done by Anthony Gregorc in the development and application of his Learning Style Delineator (Thompson & O'Brien, 1991) proposes that individuals' thinking patterns and learning styles can be measured along two threads: abstract to concrete and sequential to random. This further results in four attributes of abstract random, abstract sequential, concrete random and concrete sequential. The problem solving strategies that emanate from these could be further classified as linear-sequential and random-divergent.

Expressed by Maier (1998:23) in another way: 'Research in student learning indicates that there are individual differences in terms of how students approach their learning. Here two sets of concepts are worth highlighting: deep versus surface learning and holistic vs. serialist learners'.
Linear-thinking (sequential) individuals feel comfortable with problem solving strategies that follow a neat, orderly process. Random thinking (divergent) individuals on the other hand prefer a less confining or rigid approach.

Possible approaches to a search problem:

- **A linear approach to solving the search problem**

The first step in applying the problem solving approach should be to truly understand what the content is that you are looking for.

In the figure the different components are given for the linear (sequential) approach to problem solving.

In a search for content on the Internet the typically sequential thinking person will follow these steps.

Figure 1: A linear approach to problem solving (compare Forcier, 1996:113)
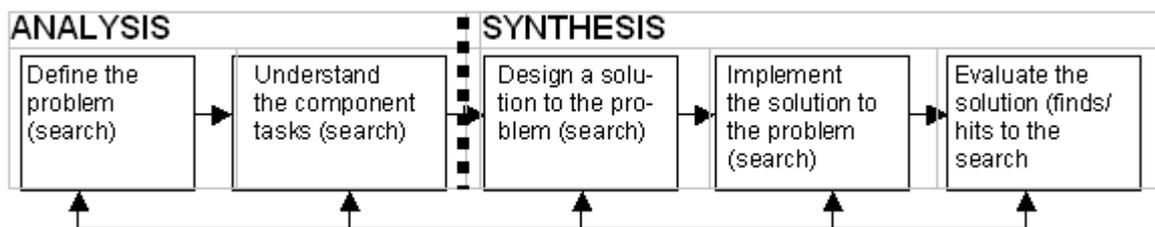
In the example given earlier the person implementing this approach will take care to first analyse and plan before venturing it into the search.

- **A non-linear approach to solving the search problem**

A person with a random or non-linear approach will look at the search problem as a whole and then start working on the fact that something must be found. Basically there are three foci in this message: Given (what is already available or in my possession), secondly, the question is, what must be found and then the procedures are decided.

Figure 2 A non-linear approach to solving a search problem
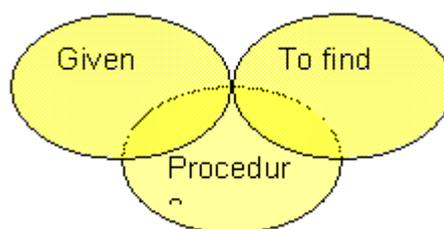
(compare Forcier, 1996:111)

2.2 Chaos theory approach

Chaos theory is a young science, and has rocketed from its obscure roots in the seventies to become one of the most fascinating. 'Chaos theory is not about disorder' (Mayer-Kress, http://www.ccsr.uiuc.edu/ gmk/). Formally, chaos theory is defined as the study of complex non-linear dynamic systems.

For purposes of this paper the only principle of this theory that we would like to propose is the fact that the chaos theory dictates that minor changes can cause huge fluctuations.

The implications of this principle for a search problem on the Internet can be illustrated as follows: If information is needed on return on equity the words equity return may be typed into the search area of a search engine. This could result in 39 542 124 finds.

Figure 3 Portion of search result based on 'equity return' (refer http://infoseek.go.com/)

In keeping with the one principle of the chaos theory that we proposed earlier, a minor change must now be made to the search string that will result in a huge 'fluctuation'. By simply adding inverted commas to the search string i.e. 'equity return' the finds are 676.
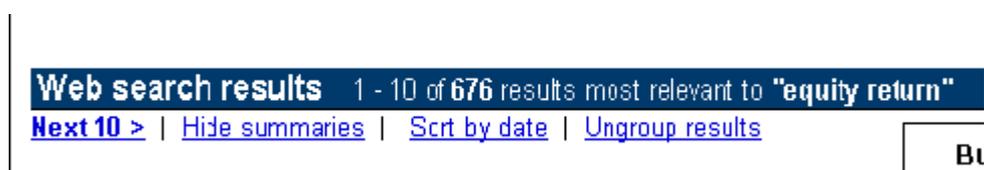
2.3 Pattern recognition

Patterns exist throughout the universe. Even seemingly chaotic systems reveal patterns normally hidden by masses of data. This statement not only describes the situation within business organisations as proposed by the authors (KKBB Management Systems), but could very well also describe the Internet if perceived as a whole. However, as with the chaos theory approach to a search problem, an analysis of data could be approached by searching for patterns.

The manager or business administrator or consultant may 'gut-feel' the solution to a lack of knowledge in a certain job related area but can start a search on the Internet and continue until a clear pattern evolves.

The 'equity return' could serve as an example again. A consultant who realises that the investments made do not yield the return could start searching for successful investors as a first step. This search could result in a hit that shows that investment in IT companies rendered the best return. The consultant may have to search further to establish a pattern which reveals the exact type of investments made or may have to start searching for information on the IT industry. The searcher will have to construct the searches according to the patterns that evolved.

Related fields are data mining approaches and principles and software that could help with the uncovering of patterns in masses of data but fall outside of this paper because here we speak of searches on the Internet to find information relevant to the self-generated curriculum of the business person and not of the analysis of data within an organisation as such.

2.4 Search engine protocol

Lastly, the search strategy that could help the searcher to find relevant hits on the Internet is the extensive use of the search protocol of the search engine being used (http://www.ouc.bc.ca/libr/connect96/search.htm).

The previous three strategies should serve as a frame of reference for the strategy now under discussion.

Every search engine has built in search facilities. Some are more sophisticated than others, but the facilities of the different search engines should be explored in order to find the power of the search facilities. Knowledge about the specific search engine and its protocol will explain why different search engines produce different data/information.

In the table that follows a comparison of a number of search engines are made.

Search engines

The many different search engines available on the WWW, and the Internet, can usually be classified as one of 5 types (compare Oliver, 1998:42; http://www.ouc.bc.ca/libr/connect96/search.htm).

TYPE	DESCRIPTION	EXAMPLE
Index/Database	Search engines (actually directories) of this form use databases of URLs together with abstracts of their contents. A search involves scanning the information in the database and returning URLs that match the search conditions. This form of search engine is comprised of URLs which have been collected and indexed manually (not automatically by a computer program).	Yahoo
Spider/Robot	This type of search engine uses a database that has been created by a program that continually and automatically scans the WWW and collects and indexes URLs.	Lycos Infoseek Magellan
Retrieval	This form of search engine takes the request from the user and scans the WWW for URLs that match. There is no database involved and each request involves a true WWW search.	Web Crawler
Harvester	A harvester is a search program that scans all the documents at sites for inclusion in a database. This sort of engine attempts to index every document in existence.	Veronica Harvester Broker
Meta-Search	An engine which uses several different engines at the same time when searching	All4One

Regardless of the search tool being used, the development of an effective search strategy is essential if you hope to obtain satisfactory results. A simplified, generic search strategy might consist of the following steps:

1. Formulate the research question and its scope.
2. Identify the important concepts within the question.
3. Identify search terms to describe those concepts.
4. Consider synonyms and variations of those terms.

5. Prepare your search logic.

This strategy should be applied to a search of any electronic information tool, including library catalogues and CD-ROM databases. However, a well-planned search strategy is of especially great importance when the database under consideration is one as large, amorphous and evolving as the World Wide Web.

When preparing the search logic an important factor to consider is the so-called Boolean operators (teachtec@calumet.yorku.ca).

Many of the search engines on the Internet offer Boolean searches. For those people unfamiliar with Boolean searches, basically what it does is help refine your search so you can get better results.

- Searching for multiple words: Boolean AND
  Looking for documents that contain words for all the colours of the British, American, and French flags? Use the Boolean AND operator:
  red AND white AND blue
  Search terms combined with the Boolean AND give back hits that contain all the terms in the record. Use it when you have dissimilar terms and want to narrow the results to a few precise hits. However, be warned that searches for John AND Hart could turn up more than references to people called John Hart - all documents with references to Gina St. John and Richard Hart would also appear.
- Searching for concepts with multiple names: Boolean OR
  When you want to search for terms with a couple of names - or for variant spellings like email and e-mail - use the Boolean OR. If you want to dig up all the email programs in shareware.com, for example, you'd search for:
  email OR e-mail
  Terms combined with OR bring up anything that matches one of your terms. This is the default operator for most search engines. It increases your total number of hits, and though you may get some irrelevancies, it's a good operator to use.
- Excluding words: Boolean NOT
  Documents found cannot contain the word after the term NOT. For example, to find documents that have the word pets, but not the word dogs, you could enter:
  pets NOT dogs
  You may also see NOT written as AND NOT or BUT NOT. This is depends on the search engine. Usually it is written as NOT.

Examples of a number of search engine home pages and interesting areas:

Infoseek
On this web page (which is the homepage of Infoseek) several of the search protocol components are revealed already. There is a choice of possible areas, such as web sites, companies, news and topics. Further there are areas created as hyperlinks where the mouse pointer changes into a hand and you could therefore 'grab' the information.
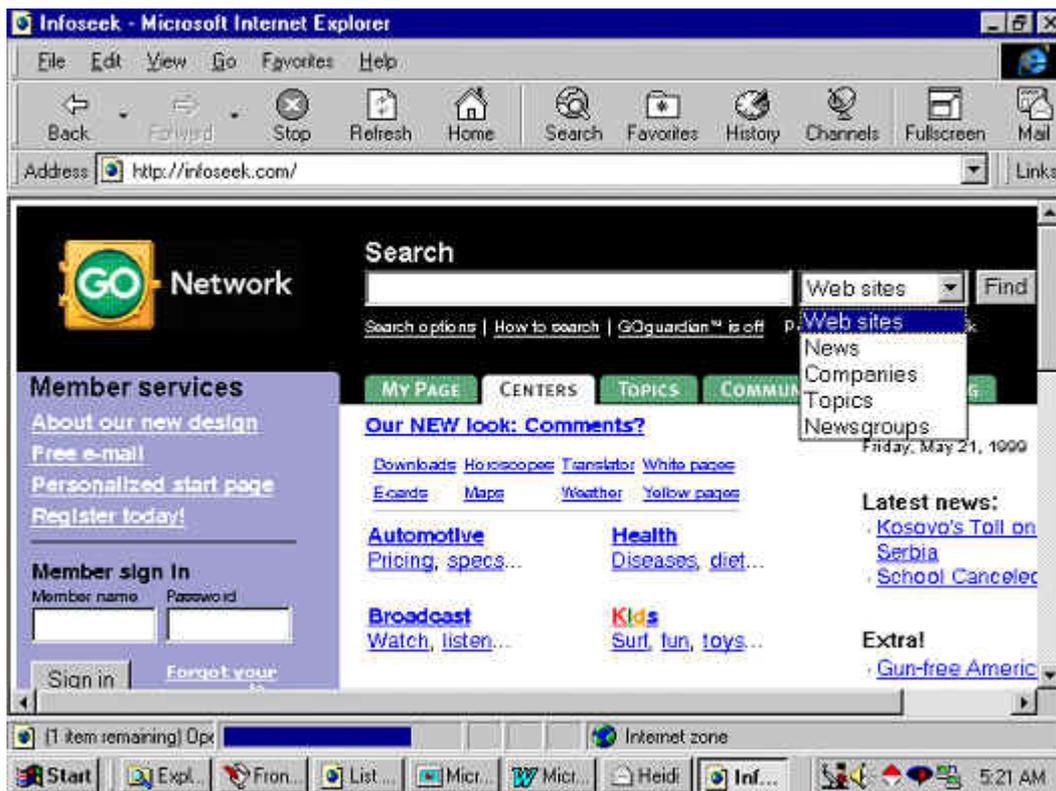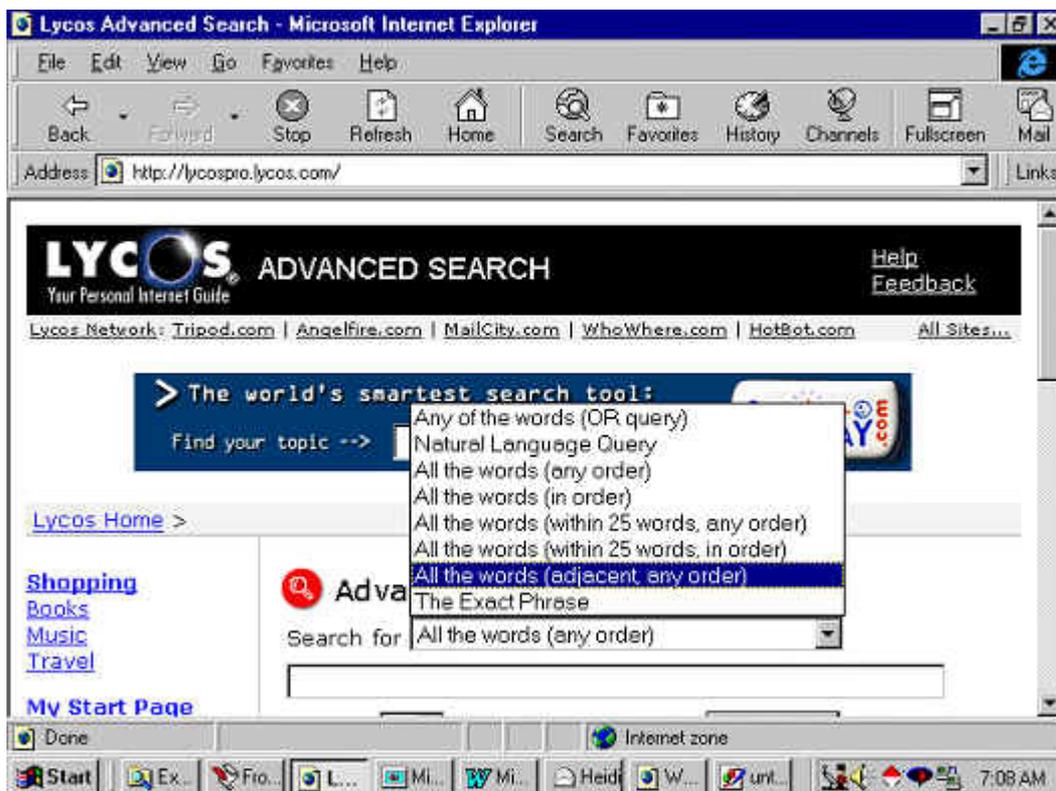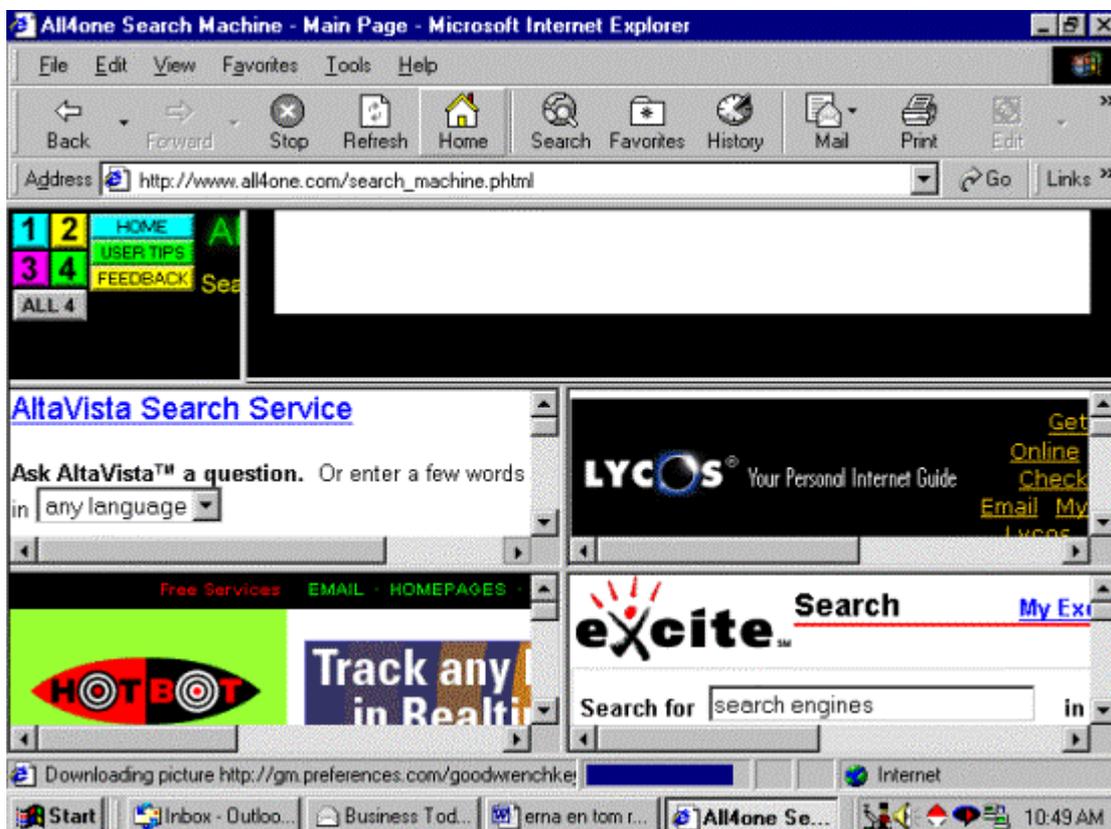
Lycos
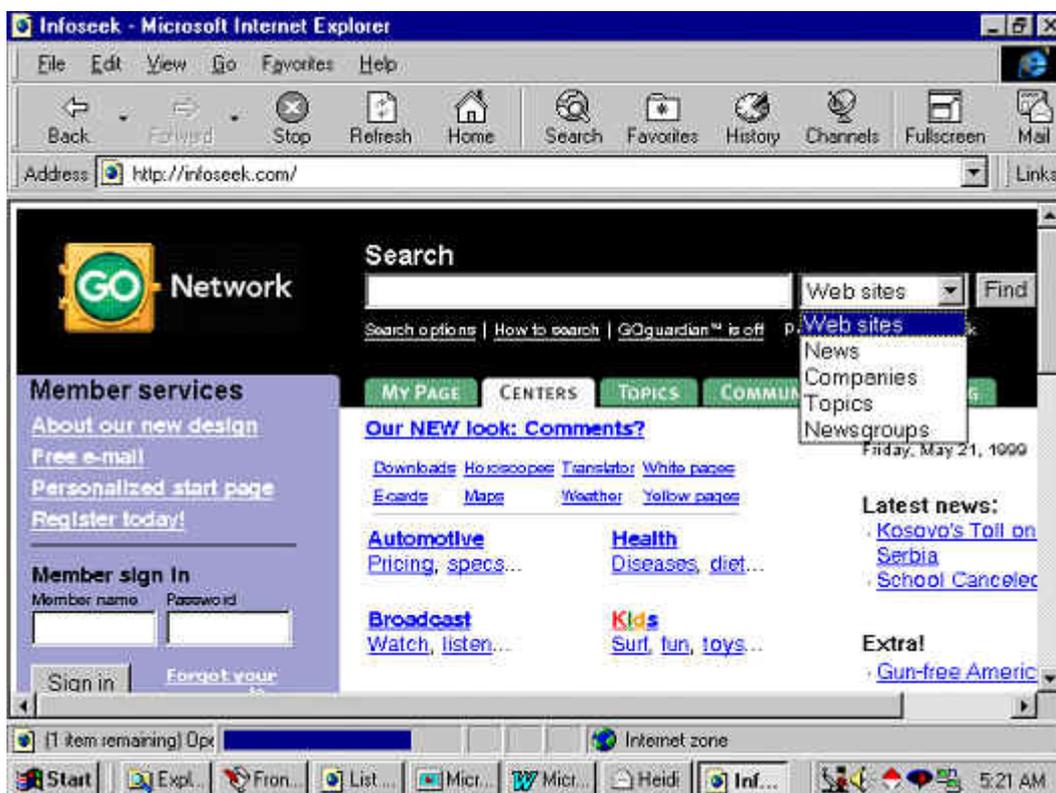
In the example of Lycos homepage above it is clear that advanced searches are possible.

The pull-down menu displayed speaks for itself and should be utilised to make the search more specified.

All4one

Four search engines display the results of a search simultaneously. The example speaks for itself.

Further discussion of specific search engine characteristics and search approaches and structures will not be included in this paper because there are multitudes of search engines each with its specific search protocol. However, these different search protocols are greatly

ignored or neglected by users of the Internet and we would like to urge those who wish to find the answers to a need for information on the Internet to utilise all available resources to access such information.

.

Conclusion

Different people have different needs. In our paper we have proposed that people need information. In the world of work, there are demands that, more often than not, cannot wait for the individual to follow a lengthy course, but who must get help in dealing with contemporary work related issues and needs on a daily basis. In order to cope with these demands the Internet could be a source of information but it must be accessed in a way that it will, indeed, be a help rather than a waste of time.

The proposed strategies could offer a possible solution to the search.

.

Further research

The above-mentioned is an exploratory attempt and your comments will be appreciated and integrated in further research.

.

References

Campbell, D. & M. 1995. Doing research on the Internet. Addison-Wesley Publishing Company, New York.

Burke, T E 1983. The philosophy of Popper. Manchester University Press, Manchester. Forcier, R.C. 1996. The computer as a productivity tool. Merrill, Englewood Cliffs, New Jersey. Columbus, Ohio.

Forsyth, I. 1998. Teaching and learning materials and the Internet. Second Edition. Kogan Page, London.

Keeves, J.P. 1996. Introduction: towards a unified view of educational research In: Keeves, J.P. (ed.). Educational research, methodology, and measurement: An international handbook. Second Edition. Cambridge University Press, Cambridge, UK.

Kent, T.W. & Mcnergney, R.F. 1999. Will technology really change education. Corwin Press, Inc. California.

KKBB management services. kkbb@hot.co.za

Maier, P., Bamett, L., Warren, A. & Brunner, D. 1998. Using technology in teaching and learning. Second Edition. Kogan Page, London.

Mayer-Kress, G. Center for Complex Systems Research Beckman Institute and Department of Physics University of Illinois at Urbana-Champaign Urbana, IL 61801 gmk@pegasos.ccsr.uiuc.edu http://www.ccsr.uiuc.edu/ gmk/

Newby, T.J., Stepich, D.A., Lehman, J.D. & Russell, J.D. 1996. Instructional technology for teaching and learning. Merrill, Columbus, Ohio.

O'Hear, A. 1980. Karl Popper. Routledge & Kegan Paul, London.

Oliva, P.F. 1992. Developing the curriculum. 3rd Edition. Harper Collins Publishers, New York.

Oliver, R. 1998. Teaching and learning on the World Wide Web. Edith Cowan University.

Popper, K. 1992. In search of a better world. Routledge, London.

Pratt, D. 1994. Curriculum planning. A handbook for professionals. Harcourt Brace College Publishers. Toronto, Canada.

Simonson, M.R. & Thompson, A. 1997. Educational computing foundations. Third Edition. Merrill, Columbus, Ohio.

Somekh, B. & Davis, N. (eds.) 1997. Using information technology effective in Teaching and Learning. Routledge, London.

Stenhouse, L. 1975. An introduction to curriculum research and development. Heinemann, London.

Internet and e-mail references

- gmk@pegasos.ccsr.uiuc.edu
- http://www.ccsr.uiuc.edu/gmk/
- http://www.ouc.bc.ca/libr/connect96/search.htm

Disclaimer